

Appl. No.: 09/418,418  
Amdt. dated December 2, 2003  
Reply to Office action of October 2, 2003

#### REMARKS/ARGUMENTS

Claims 1-12 and 14-21 are currently pending in this case. Claims 1-12 and 14-21 have been finally rejected on various grounds as discussed below. By the present amendment, Claims 1, 20 and 21 have been amended and new Claims 22 and 23 are presented. The Applicants respectfully traverse the rejections as they apply to the amended claims. Reconsideration of Claims 1-12 and 14-23, as amended, is respectfully requested in view of the following remarks.

In the response filed July 18, 2003, the Applicants' argued that the present invention is patentably distinct from the Chakrabarti reference because the reference does not form a set of expert documents from the set of all documents crawled without reference to the topic of the search query. The Applicants believed that the first element of each of Claims 1, 20 and 21 called for forming the set of experts without reference to the search query. The Examiner interpreted the claim language more broadly to include forming the expert set based on the search criteria, that is, the topic or topics of the search query.

By the present amendment, each of Claims 1, 20 and 21 has been amended to explicitly include the limitation that the set of expert documents is selected without reference to the search query. In view of these amendments, the Applicants submit that Claims 1, 20 and 21 are allowable over the Chakrabarti reference.

By the present amendment, the limitations of the original Claim 13 has been deleted from each of Claims 1, 20 and 21. A new Claim 22, intended to be an exact replacement for the original Claim 13 has been presented. In the first response in this Application, the limitations of Claim 13 had been added to Claims 1, 20 and 21, because the Examiner had indicated that Claim 13 would be allowable if rewritten in independent form. The Applicants submit that with the clarification which has now been made to Claims 1, 20 and 21, the claims are allowable over the Chakrabarti reference without the limitations of the original Claim 13. This amendment returns to the Claims to their original form except for

**Appl. No.: 09/418,418**  
**Amdt. dated December 2, 2003**  
**Reply to Office action of October 2, 2003**

addition of the express clarification that the set of expert documents is selected without reference to a search query.

By the present amendment, a new claim 23 has been presented. Claim 23 provides that the first element of Claim 1, i.e. determining which of the hypertext documents are expert documents, may occur before a search query is received. The ability to perform this part of the process before a search query is received is one of the advantages of the present invention. This preprocessing step can be performed at a convenient time, e.g. after midnight, when the computer systems may be less busy. Having this part of the process completed when a search query is received speeds up the process of performing the desired search. This advantage is in addition to the advantage that the set of expert documents is much smaller than all documents which would otherwise be searched, which also speeds up the process.

The independent Claims 1, 20 and 21 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over a publication by Chakrabarti et al. (Automatic Resource Compilation by Analyzing Hyperlink Structure and Associated Text, April 14, 1998).

The Examiner asserts that Chakrabarti teaches the step of "determining which of the hypertext documents are expert documents", which, among other limitations, is required in Claims 1, 20 and 21. The Applicants submit that Chakrabarti does not form a set of expert documents from the set of all documents crawled. Instead, each step of Chakrabarti ranks documents based on the topic of a query and forms sets of documents related to the topic. In contrast, the claimed invention requires identifying expert documents from the entire set of documents without limitation by a particular topic. Only after the claimed invention has identified the set of experts, does it use a query to identify experts relevant to a particular topic from the set created in the first step and expands the list to include target documents to which the relevant experts point.

In the present specification at page 4, lines 4-7, it is stated that: "Expert pages are preferably identified in a pre-processing step in which a subset of the pages crawled by the search engine are identified as experts (for example, 2.5

**Appl. No.: 09/418,418**  
**Amdt. dated December 2, 2003**  
**Reply to Office action of October 2, 2003**

million of 140 million pages may be identified as experts)." Such a pre-processing step may be done for all pages crawled by the search engine, e.g. 140 million pages. It may be done before a query is actually received, since it does not use a topic in the first step. The set of experts is not limited to any particular topic submitted in any particular query.

In describing Fig. 2(a), at page 6, lines 23-27 the present specification states that: "...an initial set of hypertext pages 202 is obtained by, for example, a crawl of the world wide web. The hypertext documents (e.g. pages) are processed to yield a set of expert documents 204. When a query is received, the expert documents are ranked in accordance with the query to yield ranked expert documents 206." This is also stated on Page 8, lines 25-27 as follows: "Step 202 preferably processes a search engine's database of pages and selects a subset which is considered to be good sources of links on specific topics, albeit unknown."

The expert list is created by measuring the out-degree of each page and checking to be sure the page points to non-affiliated hosts, see for example Fig. 3(a). In Fig. 3(b) the experts may be further limited to those which point to pages with the same broad classification, but this is not limited to any particular topic. Since the expert set creation step does not depend on having received a query with a topic, it can be performed occasionally, for example at off peak demand times, and stored for use when queries are received.

The second step of the present claims is to rank the expert documents in accordance with a current search query. This is the first step in which the topic of the query becomes part of the process. The preselected expert list may be searched for pages relevant to the topic, not the entire set of pages. That is, it is not necessary to search the entire set of pages crawled by the search engine. In the above example, only 2.5 million pages needs to be searched, not 140 million. The present invention then uses the ranked expert pages which are relevant to the topic to expand the list of target documents by ranking the pages to which the relevant experts point. It is not necessary to search the entire web to find

**Appl. No.: 09/418,418**  
**Amdt. dated December 2, 2003**  
**Reply to Office action of October 2, 2003**

documents relevant to the topic, because the ranked expert documents point to them.

In contrast, the process of Chakrabarti begins with a topic from a query and searches the entire set of pages accessible to the search engine for pages relevant to the topic. For example in Section 2. Algorithm, second paragraph, Chakrabarti states: "Given a topic, the algorithm first gathers a collection of pages from among which it will distill ones that it considers the best for the topic. ... The topic is sent to a term-based search engine – *AltaVista* in our case – and a root set of 200 documents containing the topic term(s) is collected. ... The root set is then augmented through the following *expansion* step: we add to the root set (1) any document that points to a document in the root set, and (2) any document that is pointed to by a document in the root set. We perform this expansion step twice..." To the extent that Chakrabarti expands by identifying documents that point to the root set, it is searching the entire web again.

Chakrabarti then uses hub scores and authority scores to rank the documents in the augmented set. In Section 2. Algorithm, fourth paragraph, Chakrabarti discusses generation of a hub score  $h(p)$  and an authority score  $a(p)$  for each page in the augmented set. In Section 2. Algorithm, fifth paragraph, Chakrabarti discusses weighting the hub and authority scores based on topic references as explained in more detail in Section 2.1. In Section 2. Algorithm, sixth paragraph, Chakrabarti computes vectors of the weighted hub and authority scores and selects the pages with the top fifteen scores for each of hub and authority values. Chakrabarti does not actually discuss experts, much less identifying a set of experts without reference to a particular topic submitted in a query.

In view of the above remarks, the Applicants submit that Claims 1, 20 and 21 are clearly patentable over the Chakrabarti reference. Since the remaining claims all depend from Claims 1, 20 and 21, the remaining claims should also be patentable over the Chakrabarti reference. Applicants respectfully request reconsideration and allowance of Claims 1-12 and 14-23.

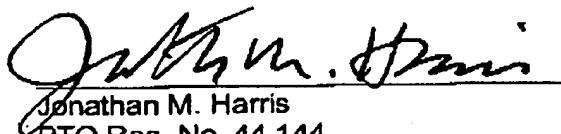
Appl. No.: 09/418,418  
Amdt. dated December 2, 2003  
Reply to Office action of October 2, 2003

If the Examiner feels that a telephone conference would expedite the resolution of this case, he is respectfully requested to contact the undersigned.

In the course of the foregoing discussions, Applicants may have at times referred to claim limitations in shorthand fashion, or may have focused on a particular claim element. This discussion should not be interpreted to mean that the other limitations can be ignored or dismissed. The claims must be viewed as a whole, and each limitation of the claims must be considered when determining the patentability of the claims. Moreover, it should be understood that there may be other distinctions between the claims and the prior art which have yet to be raised, but which may be raised in the future.

If any fees or time extensions are inadvertently omitted or if any fees have been overpaid, please appropriately charge or credit those fees to Hewlett-Packard Company Deposit Account Number 08-2025 and enter any time extension(s) necessary to prevent this case from being abandoned.

Respectfully submitted,



Jonathan M. Harris  
PTO Reg. No. 44,144  
CONLEY ROSE, P.C.  
(713) 238-8000 (Phone)  
(713) 238-8008 (Fax)  
ATTORNEY FOR APPLICANTS

HEWLETT-PACKARD COMPANY  
Intellectual Property Administration  
Legal Dept., M/S 35  
P.O. Box 272400  
Fort Collins, CO 80527-2400